A posteriori error estimate in fluid simulations of turbulent edge plasmas for magnetic fusion in tokamak using the data mining iPoPe method

T. Cartier-Michaud, D. Galassi, Ph Ghendrih, P. Tamain, F. Schwander, E. Serre

A posteriori error estimate in fluid simulations of turbulent edge plasmas for magnetic fusion in tokamak using the data mining iPoPe method

T. Cartier-Michaud,[1, a)] D. Galassi,[1] Ph. Ghendrih,[2] P. Tamain,[2] F. Schwander,[1] and E. Serre[1]
[1)] *Aix-Marseille Univ., CNRS, Centrale Marseille, M2P2, France*
[2)] *CEA, IRFM, F-13108 Saint-Paul-Lez-Durance, France*

(Dated: 20 May 2020)

Progressing towards more reliable numerical solutions in the simulation of plasma for magnetic confinement fusion has become a critical issue for the success of the ITER operation. This requires developing rigorous and efficient methods of verification of the numerical simulations in any relevant flow regimes of the operation. The paper introduces a new formulation of the PoPe[1] method, namely the independent Projection on Proper elements method (iPoPe) to quantify the numerical error by performing a data-driven identification of the mathematical model from the simulation outputs. Based on a statistical postprocessing of the outputs database, the method provides a measure of the error by estimating the distance between the (numerical) effective and (analytical) theoretical weights of each operator implemented in the mathematical model. The efficiency of the present method is illustrated on turbulent edge plasma simulations based on a drift-reduced Braginskii fluid model in realistic magnetic geometries. Results show the effective order of the numerical method in these multiscale flow regimes as well as the values of the plasma parameters which can be safely simulated with respect to a given discretization. In this sense, the method goes one step further than the Method of Manufactured Solution (MMS[2–4]), recently introduced in fusion, and provides an efficient verification procedure of the numerical simulations in any regimes, including turbulent ones that could be generalized to other scientific domains.

PACS numbers: Valid PACS appear here
Keywords: PoPe, verification method, edge plasma modelling, TOKAM3X

I. INTRODUCTION

Fusion based on magnetic confinement aims at producing carbon-free electric power by using the energy liberated by fusing deuterium and tritium nuclei at extremely high temperatures (10^7-10^8K), within a plasma confined by magnetic fields in machines of toroidal shape known as tokamaks. The success of future fusion experiments in reactors like ITER is conditioned by our capability to both ensure the quality of plasma confinement in the core, and to control the heat and particle fluxes on the wall. During the operation, optimized plasma scenarii will have to be found to satisfy conflicting constraints between the fusion ignition that requires the plasma to be sufficiently heated, fueled and confined, and the heat exhaust on the tokamak wall to avoid undue damage. Such scenarii still remain uncertain, largely owing to an incomplete understanding of the mechanisms at play.

The difficulty to generate a global description based on experimental measurements in tokamaks requires performing complementary numerical simulations. With the increasing performance of numerical codes and of the computational resources, such simulations are becoming a competitive way to investigate fusion plasmas physics as well as to design optimized plasma scenarii during tokamak operation. However, their reliability is still acknowledged by the international community as being a critical issue whose a rigorous assessment can be carried out by using the Verification & Validation procedure (V & V[5]). The verification procedure aims to check the correct implementation of the mathematical model in the code whereas the validation procedure checks the ability of the numerical solutions to describe the true physics.

Developed by the computational fluid dynamics community, the Method of the Manufactured solutions MMS (see e.g. Ref[2–4]) has been recently introduced in fusion to verify the implementation of a discretized solution procedure[6,7]. In this method, a manufactured analytical function is imposed as the solution of the problem by adding a corrective source term into the mathematical model. The method then works very similarly to a Richardson Extrapolation[8] based on mesh and time step sensitivity study. The error between the numerical solution provided by the code and the imposed manufactured solution is screened when varying mesh and time step sizes. The method is generally applied on smooth and rather simple solutions in laminar regimes to be easily derived. In addition, the method must ensure that the error clearly decreases towards zero as the discretization becomes small and that the theoretical order of convergence can be attained. However, these analytical solutions are clearly not representative of production runs.

The Projection on Proper elements method (PoPe) has been introduced in Ref.[1] and allows data-driven fluid and kinetic model identification using simulations outputs. Investigations have been conducted in Ref.[9], in which different noises were added to the simulations outputs to mimic noise in diagnosis of real experiments. Studies

[a)] Electronic mail: t.cartiermichaud@gmail.com

have not yet been pursued in that direction as the main aim of developing (i)PoPe was to perform codes verification and models reduction. The estimation of the error in the simulation relies on a statistical measure of the distance between (numerical) effective and (analytical) theoretical weights of the various operators implemented in the mathematical model to solve. Unlike the MMS, it is enabled to address any flow regimes, including turbulence, and in that sense goes one step further than the MMS classically used.

Regarding the investigation and simulation of the edge turbulence, despite the exponential growth of computer speed along with significant improvements in computer technology, kinetic simulations remain extremely costly from the computational point of view, even if pioneering full-f gyrokinetic simulations of the edge start appearing in the fusion community, addressing physical phenomena of fundamental interest for fusion operation like transport barrier formation[10,11]. As a consequence, fluid approach based on Braginskii equations[12] and drift ordering[13] remains a standard one near the wall where the temperature is lower and the collisional mean free path significantly smaller than in the core. Various 3D codes already exist in the community, generally based on first and second-order finite-differences / finite-volumes numerical schemes (see a recent review of most of these codes in Refs.[7]). Let's mention BOUT++ [US, UK][14], GBS [CH][15], CYTO [Dan, UK][16] and TOKAM3X[17][FR] for the most commonly used in the international community. In this paper, we will consider TOKAM3X, which is co-developed in our team. TOKAM3X has been formerly verified on smooth solutions using the MMS[17]. The MMS study had to be performed on one equation at a time, using smooth solutions in space and time with respect to the discretization (e.g. $\propto cos(2\pi t)\sin(2\pi r/a)\sin(\theta)\sin(\varphi)$). Moreover, the verification method was applied exclusively in a simple geometry (slab or limiter with circular cross section), and by prescribing simplified boundary conditions (usually Dirichlet's). It is thus quite far from the turbulent regimes targeted by the code, and under these conditions, some operators might not be involved in the solution, leading to an incomplete verification of the equations. However, the convergence rates of the whole numerical scheme was recovered with this procedure, excepted a slight deviation on the vorticity equation. In addition, TOKAM3X was benchmarked with other codes of the fusion community by simulating seeded blobs dynamics in simple geometries[18]. It has been also punctually validated with respect to plasma equilibria measured in TORE-SUPRA at various limiter positions and known as MISTRAL test case[19]. Three-dimensional turbulent fluid solutions have been recently investigated for various magnetic equilibria in limited[6] and for the first time in diverted geometry[20], and have shown the capacity of TOKAM3X to provide original results in qualitative agreement with respect to experimental measurements and theoretical expectations. However, a rigorous point-by-point validation with experiments remains unreachable due to resolution limits, to the incompleteness of the physical model, and the error margins of experimental measurements.

In this paper, we investigate the numerical error using a new release of PoPe, named iPoPe, on two turbulent solutions provided by TOKAM3X in both limited and diverted magnetic geometries. The physical parameters values, the mesh size and the geometries make these solutions significant of production solutions usually considered with this code. A brief overview of the TOKAM3X numerical ingredients is first provided in Sec. II. The new formulation of the PoPe method previously detailed in Ref.[1], named the independent PoPe (iPoPe) method together with the associated verification procedure are provided in Sec. III. The main features of the two 3D turbulent solutions analysis are then introduced in Sec. IV. A first study of iPoPe's weights and residuals is shown in Sec. V, before doing a scan onto the space resolution in Sec. VI as well as on the perpendicular diffusion coefficient in Sec. VII. The impact of the time discretization is out of the scope of this paper but final concluding remarks and discussions are provided in Sec. VIII.

II. A SHORT OVERVIEW OF THE TOKAM3X FLUID MODEL

TOKAM3X provides a first principle description of cross-field turbulence that allows to simulate turbulence self-consistently without any scale separation between fluctuations and geometrical scales of the device. The code is based on a drift-reduced Braginskii model, associated to a set of boundary conditions (Bohm conditions in the parallel condition) at the target plates. All details concerning the mathematical model and numerical schemes can be found in Ref.[17].

A. The mathematical model

Under some hypothesis and ordering detailed in Ref.[17], four dimensionless conservation equations are derived for the electronic density N, the ionic parallel momentum Γ, the electrostatic potential Φ and the parallel current $j_\parallel$ which defines the parallel advection velocity for electrons. They write as:

$$\begin{cases} \partial_t N + \vec{\nabla}\cdot\left(N\vec{u}^e\right) = S_N + \vec{\nabla}\cdot\left(D_N\vec{\nabla}_\perp N\right) \\ \partial_t \Gamma + \vec{\nabla}\cdot\left(\Gamma\vec{u}^i\right) = -\nabla_\parallel P + \vec{\nabla}\cdot\left(D_\Gamma\vec{\nabla}_\perp \Gamma\right) \\ \eta_\parallel N j_\parallel = -N\nabla_\parallel\phi + \nabla_\parallel N \\ \nabla.\vec{j} = 0 \end{cases} \quad (1)$$

where S_N is a volumetric source term driving the particle flux, where $\parallel$ and $\perp$ define the direction parallel and perpendicular to the magnetic field $\vec{b} = \vec{B}/||\vec{B}||$ with $\vec{u}^{e,i} = u_\parallel^{e,i}\vec{b} + \vec{u}_\perp^{e,i}$, and where $\eta_\parallel$ is the normalized parallel collisional resistivity of the plasma. The last equation

corresponds to the charge balance when using the quasi-neutrality assumption ($\nabla.\vec{j} = 0$). It is treated through the vorticity W such that:

$$\begin{cases} \partial_t W + \vec{\nabla}\cdot(W\vec{u}^i) = \vec{\nabla}\cdot\left(N\left(\vec{u}^i_{\nabla B} - \vec{u}^e_{\nabla B}\right) + j_\parallel \vec{b}\right) + \\ \vec{\nabla}\cdot(D_W\vec{\nabla}_\perp W) \\ W = \vec{\nabla}\cdot\left(\frac{1}{B^2}(\vec{\nabla}_\perp\phi + \frac{1}{N}\vec{\nabla}_\perp N)\right) \end{cases} \quad (2)$$

In all equations, effective diffusion in the cross-field direction account for collisional transport, and it crudely models the impact of turbulent small scales. Diffusion in the parallel direction has been neglected with respect to the parallel convection. $D_{N,\Gamma,W}$ are kept constants in the present model, and are usually equal or smaller than 0.1.

Boundary conditions are defined as follows:

- At the core and at the external wall and in the perpendicular direction, homogeneous Neumann conditions are imposed for all variables, $\partial_\perp(\cdot) = 0$.
- In the Scrape-Off Layer (SOL, where magnetic field lines intercept the target plates (limiter or divertor)) and in the parallel direction, usual Bohm boundary conditions are used[21] . They model the physics of the sheath located next to the limiter wall and in which quasi-neutrality is no longer valid. An isothermal sonic velocity $u \geq \pm c_s$ and the parallel current $j_\parallel$ are imposed on the two opposite sides of the target plates that leads in the dimensionless form to:

$$|\Gamma| \geq N \quad (3)$$

$$j_\parallel = \pm N(1 - exp(\Lambda - \phi)) \simeq \pm N(\Lambda - \phi) \quad (4)$$

where Λ is the sheath floating potential.

The condition on the parallel gradient of the potential used to solve Eq. 1 is derived from Eq. 4 and using the generalized Ohm's law (Eq. 1) and writes as follows:

$$\nabla_\parallel\phi = \pm\eta_\parallel N(\Lambda - \phi) + \frac{\nabla_\parallel N}{N} \quad (5)$$

For N, the second normal derivative is imposed to zero.

B. Numerical details

The equations shown above correspond to those of a compressible adiabatic gas flow in the parallel direction (fast scales). In the perpendicular direction, they correspond to those of an incompressible flow, dominated by turbulent processes. Thus, the algorithm has been constructed to split the discretization of the parallel and perpendicular directions.

The numerical approximation is based on an advanced second-order accurate finite-differences/finite-volumes method. It is associated to a 2nd-order WENO reconstruction used for the advection terms when dealing with shocks. The code is multithreated using a hybrid OpenMP/MPI parallelization. A multidomain decomposition allows us to map the physical space into a set of rectangular subdomains keeping a structured flux-surfaces aligned mesh in any versatile geometries. For the time evolution, a semi-implicit scheme based on a first-order operator splitting is used. Such a scheme is adapted for advancing the parallel current terms associated to an extremely fast dynamics and thus requiring an implicit treatment. In the charge balance equation, the three spatial directions are coupled. As a consequence, the vorticity operator to invert is fully $3D$ and badly conditioned due to small values of the parallel resistivity in tokamak plasma ($\eta_\parallel \approx 10^{-5} - 10^{-8}$, normalized values). This particularity has till now hindered the use of an efficient iterative scheme. Being time-independent, this operator is currently inverted before the time loop using a direct method (LU decomposition) included in the PASTIX library[22].

III. THE IPOPE METHOD IN A NUTSHELL: DECOMPOSITION ONTO A RELEVANT BASIS PLUS AN ERROR

The independent Projection on Proper elements (iPoPe) method has been extended from the PoPe method recently detailed in Ref.[1]. Based on statistics done on the outputs data of the code, these methods enable to provide an estimate of the numerical error carried out by simulations in any flow regimes by measuring the distance between the theoretical set of equations and the effective set extracted from the simulation outputs. In this sense they can be interpreted both as data-driven model identification tools and as powerful tools for codes verification. To describe the method in a general way an advection-diffusion equation of any scalar variable Ψ, shaped to be significant of turbulence fluid codes in fusion, is considered:

$$\partial_t\Psi = \vec{\nabla}\cdot(D_\Psi\vec{\nabla}_\perp\Psi) - \vec{\nabla}\cdot(\Psi\vec{u}) + S_\Psi \quad (6)$$

where $\vec{u}$ is the advection velocity parallel to the magnetic field line, D_Ψ is the diffusion coefficient in the perpendicular direction and S_Ψ is a source term. We can immediately rewrite Eq.(6) in a more compact form using the Einstein's convention on indexes.

$$\begin{aligned} \{\partial_t\Psi\}_{th} &= w^i_{th}\{O^i\}_{th} \\ \{O^i\}_{th} &= [\vec{\nabla}\cdot(D_\Psi\vec{\nabla}_\perp\Psi),\ -\vec{\nabla}\cdot(\Psi\vec{u}),\ S_\Psi] \end{aligned} \quad (7)$$

For $i = 1, 2, 3$, $\{O^i\}_{th}$ names the three operators on the right hand side of Eq.(6) and using the three associated weights $w^i_{th} = [+1, +1, +1]$.

In the following, notations "$\{.\}_{th}$", "$\{.\}_{ef}$", "$\{.\}_{ol}$" stand for "theoretical", "effective" and "off-line" evaluation of the operators, respectively.

A. The iPoPe procedure

Three steps are actually required :

1. The identification and measurement of the actual unknowns of the problem from the code outputs.
   In a time evolving equation such as Eq.(7), the main unknown is the time derivative of Ψ. Its measurements performed with finite-differences provide the "effective" time evolution of the system simulated by the simulation code, and noted $\{\partial_t\Psi\}_{ef}$. In order to avoid additional error introduced by iPoPe, a high-order finite-differences scheme is required. A fourth-order scheme Eq.(8) is used for TOKAM3X as this code used a second-order scheme and forth-order has been proved to be accurate enough with respect to the fastest time scales. Such computation can be performed iteratively, without storing the solution at different time steps between $t-2\Delta t$ and $t+2\Delta t$.

$$\{\partial_t\Psi\}_{ef}(t) = \sum_{j=-2}^{2} c(j)\Psi(t+j\Delta t) + \mathcal{O}(\Delta t^4) \qquad (8)$$
$$c(-2:2) = [+1,-8,0,+8,-1]/(12\Delta t)$$

2. The computation of each operator in the model concerned in post-processing and with high accuracy.
   Similarly to the previous point, while TOKAM3X makes use of second-order methods in space, the "off-line" estimate uses a sixth-order scheme. This sixth-order has been chosen using Fig. 2 and Fig. 3, showing that going to higher order would not change significantly the results of the computations. This is a key aspect of the verification process which writes as:

$$\|\{O^i\}_{th} - \{O^i\}_{ol}\| \ll \|\{O^i\}_{th} - \{O^i\}_{ef}\| \qquad (9)$$

   so that

$$\{O^i\}_{ol} \simeq \{O^i\}_{th} \qquad (10)$$

3. A projection method to decompose the dynamics $\{\partial_t\Psi\}_{ef}$ onto the operators $\{O^i\}_{ol}$.
   We finally project each operator independently of each other by, in the following order, computing the total residual ϵ_{to}, then projecting operators one at a time to get effective weights w^i_{ef} and finally computing the effective residual ϵ_{ef}:

$$\epsilon_{to} = \{\partial_t\Psi\}_{ef} - \{\partial_t\Psi\}_{th} \qquad (11)$$
$$= \{\partial_t\Psi\}_{ef} - \omega^i_{th}\{\mathcal{O}^i\}_{ol}$$
$$\omega^i_{ef} = \omega^i_{th} + \epsilon_{to} \times \{\mathcal{O}^i\}_{ol}/||\{\mathcal{O}^i\}_{ol}||^2 \qquad (12)$$
$$\epsilon_{ef} = \{\partial_t\Psi\}_{ef} - \omega^i_{ef}\{\mathcal{O}^i\}_{ol} \qquad (13)$$

B. iPoPe versus PoPe

Unlike for iPoPe, with the PoPe method[1] all operators were projected at once by solving the following linear system based on the least mean square algorithm:

$$A^t A\ w_{ef} = A^t\{\partial_t\Psi\}_{ef} \qquad (14)$$

where the $P \times I$ matrix A is defined as

$$A(p,i) \equiv \{O(p)^i\}_{ol} \qquad (15)$$

Each column of the matrix A is the evaluation of the i^{th} operator from the set of I operators in the tested equation. These evaluations are performed for P points labelled by the index p. P is defined by the discretization used to solve the equation with the code that we want to verify in the first place. Usually $P \gg I$, so that a large number of points can be considered in order to reduce the statistical error in the estimation of w_{ef}. For TOKAM3X, typically, we have $P = N_r \times N_\theta \times N_\varphi \times N_t \gg 10^6$ for a unique simulation. Nevertheless, in order to introduce a time dependence in w_{ef}, we usually use N_t sets of $N_r \times N_\theta \times N_\varphi \geq 10^4$ points.

The matrix A^tA is of size $I \times I$, 3×3 in the case of Eq.(6), and can be solve at very low computational cost. The residual is then recovered by explicitly computing:

$$\{\partial_t\Psi\}_{ef} - A\ w_{ef} = \epsilon_{ef} \qquad (16)$$

According to the least mean square approach, weights w^i_{ef} are determined under the constraint of minimizing $||\epsilon_{ef}||_2$. Other systems than (14) can be build to perform minimization of ϵ_{ef} under different norms and/or more constraints.

The interest of iPoPe with respect to PoPe, it is that it does not require to build the matrices A and A^tA nor the right hand side $A^t\{\partial_t\Psi\}_{ef}$. This can lead to several optimisation from a computational point of view and remove any constraint of solvability of the linear system. Indeed, to get reliable outputs from PoPe, one has to insure that this linear system is not singular or that its condition number is not too high. Thus, iPoPe is no longer based on an optimization problem. It does not minimize the $\mathcal{L}^2$-norm of the residual but directly interpret the total error ϵ_{to} on each operator as a "worst case senario". Last but not least, in the case of an equilibrium, $\{\partial_t\Psi\}_{ef} = 0$, the trivial solution of Eq.(14) is $w_{ef} = 0$, regardless of if $\vec{\nabla}\cdot(D_\Psi\vec{\nabla}_\perp\Psi) - \vec{\nabla}\cdot(\Psi\vec{u}) + S_\Psi = 0$. In such case one should either: 1) rewrite the theoretical equation,

removing term ∂_t as it has a strictly null contribution and projecting using another operator as the right hand side, or 2) search for the kernel of the same linear system. Using iPoPe instead of PoPe, this situation is avoided.

Both PoPe and iPoPe lead to the interpretation of the effective time derivative of a given code $\{\partial_t \Psi\}_{ef}$ as a sum of operators $\{O\}^i_{ol}$, each weighted by w^i_{ef}, plus a residual ϵ_{ef}. In the case of PoPe, the effective residual has no linear dependency with respect to the operators of the equations, while the effective residual of iPoPe can have linear dependencies and it is a worst case scenario in that sense. Ideally, $w_{ef} = w_{th}$ and $\epsilon_{ef} = 0$. Such weights plus residual decomposition is relevant since effective weights define the nature of the equation. They control the behaviour of the system theoretically, thus numerically. Compared to a Principal Component Analysis, this approach decomposes the unknown onto a basis of vectors with obvious physical meanings, the operators.

A way to illustrate the independency of the projection step with iPoPe is to study a simple error, such as the use of a diffusion coefficient two times higher than theoretically expected. Using PoPe, this would be directly identified with $w^1_{ef} = 2$ instead of the value $w^1_{th} = 1$, and without impacting $\epsilon_{ef} = 0$. Here with iPoPe, one would recover $w^1_{ef} = 2$ again, but other weights w^i_{ef} could be affected if $\epsilon_{to} \times \mathcal{O}^i_{ol} \neq 0$. Then the condition $\epsilon_{ef} = 0$ could be likely to vanish. When projecting the second and the third operators using iPoPe, the first one is not taken into account, despite the fact the first one would have captured the entire error. The projection of one operator is thus independent of the other operators and in that sense this is a worst case scenario of expressing the complete error on each operator. Finally the computations of each weight w^i_{ef} could even be performed in parallel as it only depends on ϵ_{to} and on the i-th operator.

Both PoPe and iPoPe descriptions, weights plus residual, are the decomposition of the same dynamic: bijections exist between 1) the actual dynamics ($\{\partial_t \Psi\}_{ef}$), 2) the PoPe decomposition and 3) the iPoPe decomposition. Still, PoPe concentrates more information in the weights as the residual is under constraint of $\mathcal{L}^2$-norm minimization while iPoPe tends to keep more information in the residual.

Analysing the weights would give absolute confidence if $\{O^i\}_{ef}$, operators effectively calculated within the code under study, were "exact". As we discretize solutions over a finite number of degrees of freedom, each of them having a finite accuracy, operators $\{O^i\}_{ef}$ usually differ from theoretical expression of operators $\{O^i\}_{th}$. The theoretical expression of operators being not usually accessible, we do not compare $\{O^i\}_{ef}$ to $\{O^i\}_{th}$ but rather to $\{O^i\}_{ol}$, a set of operators computed in post-processing using a greater accuracy than for $\{O^i\}_{ef}$, as already expressed in Eq.(9). This point is important to be able to associate the residual ϵ_{ef} to an error of the code and not to an error in the verification process. In the present case, authors considered sufficient the sixth-order estimation of operators as $||w^i_{th}\{O^i\}_{ol}|| = ||\partial_t X||$ in Fig. 2 and Fig. 3 shows a discrepancy between sixth-order and fourth-order significantly smaller compared to the discrepancy between second-order and fourth-order. An interesting point is that if $\{O^i\}_{ef} = \{O^i\}_{ol}\ \forall\ i \in [1,3]$ in Eq.(6), then necessarily $\epsilon_{ef} = 0$ and one can only study the correctness of $w_{ef} = w_{th}$ and not if the computation of operators is correct itself. Off-line greater accuracy is easily obtained and at a reasonable computational cost because for off-line computations we do not have to take care of any stability with respect to time integration.

It is interesting to note that for time explicit numerical schemes the time step is defined in order to ensure small variations of the system between two integration time steps. As the solution is generally not changing significantly between two time steps, it is not mandatory to verify each time step of the code. A reasonable frequency of verification would be the frequency used to save diagnosis in order to then investigate physics. Indeed, this is the relevant time scale for the variations of the system. For explicit numerical scheme it might be of the order of $0.1 - 5\%$ of the time steps computed by the code. For TOKAM3X, the time integration constant chosen is $\Delta t_{numeric} = 1$ while a conservative frequency of $\Delta t_{diag} = 50$ has been chosen for diagnosis saving and verification. Thus only 2% of the time integration step are verified while still being representative of the dynamic. It is worth to recall that the time derivative used by (i)PoPe is nevertheless computed using a time step of $\Delta t_{numeric} = 1$. Combining the lower cost of off-line estimation and the small fraction of time step verified, the overall computational cost is then negligible, less than 1%.

C. Definitions and (i)PoPe indexes

Using again the assumption $\{O^i\}_{ol} = \{O^i\}_{th}$, we get six useful definitions:

- $\{\partial_t \Psi\}_{th} = w^i_{th}\{O^i\}_{ol}$ is the theoretical time evolution of the system. This is the evolution theoreticians would work on analytically, including no source of errors and uncertainties induced by any numerical solver.

- $\{\partial_t \Psi\}_{ef} = w^i_{ef}\{O^i\}_{ol} + \epsilon_{ef}$ is the effective time evolution of the system simulated by the code. It includes numerical errors due to approximations by numerical methods. It could even contain possible bugs if the code was not correct.

- $\{\partial_t \Psi\}_{po} = w^i_{ef}\{O^i\}_{ol}$ is the effective time evolution of the system without taking into account ϵ_{ef}, but only defined by the effective weights. If $||\epsilon_{ef}||$ has a small impact on the dynamics but the coefficients w^i_{ef} are not close enough to w^i_{th}, then it would be the expression theoreticians should use to interpret the results instead of of $\{\partial_t \Psi\}_{th} = w^i_{th}\{O^i\}_{ol}$.

- $\epsilon_{to} = \{\partial_t \Psi\}_{ef} - \{\partial_t \Psi\}_{th}$ is the total residual which contains the full discrepancy between the theoretical and the effective evolution. As a rule of thumbs, the smaller is $||\epsilon_{to}||$, the better is the simulation. When choosing numerical methods and resolutions of the discretization, this criterion should be the first to minimize.

- $\epsilon_{ef} = \{\partial_t \Psi\}_{ef} - \{\partial_t \Psi\}_{po}$ is the effective residual which does not contain any linear dependency with respect to $\{O^i\}_{ol}$ when using PoPe, as introduced in Eq.(16). In order to take $\{\partial_t \Psi\}_{po}$ as the effective equations of the model, one has to insure ϵ_{ef} has an acceptable impact on the general dynamics. Studies of the evolution of the residual due to linear operators ($\mathcal{O}^i(\epsilon_{ef})$), coupling due to non linear operators (($\mathcal{O}^i(X, \epsilon_{ef})$) and tracking different statistical momenta ($< \epsilon_{ef}^m >_{r,\theta,\phi,t}$) are advised.

- $\epsilon_{co} = \{\partial_t \Psi\}_{po} - \{\partial_t \Psi\}_{th}$ is the residual carried by the discrepancy between w_{ef}^i and w_{th}^i, a residual that could be partially compensated by changing the parameters used to run the code if they can be simulated. One can notice that $\epsilon_{to} = \epsilon_{ef} + \epsilon_{co}$.

Finally, one can define the "(i)PoPe index" such as $I_{(i)PoPe} = -log_{10}(||\epsilon_{to}||_2 / ||\{\partial_t\}_{ef}||_2)$. Using the weights of the operators would lead to $I_{(i)PoPe} = min_i(-log_{10}(|w_{ef}^i - w_{th}^i|))$. According to these definitions, a perfect simulation would lead to $\epsilon_{to} \to 0$ and $w_{ef}^i \to w_{th}^i$ so $I_{(i)PoPe} \to \infty$, while in practice the $I_{(i)PoPe}$ index is always finite. To give a physical meaning to the values of the (i)PoPe index is however an ongoing work. Measuring the discrepancy between the theoretical and the effective values of the operators weights is easier than predicting the impact of this discrepancy on the solution. In some cases, predicting the behavior of the error could be as challenging as predicting the behavior of the system without the error. This point will be discussed again in Sec. VIII.

IV. TURBULENT SOLUTION ANALYSIS

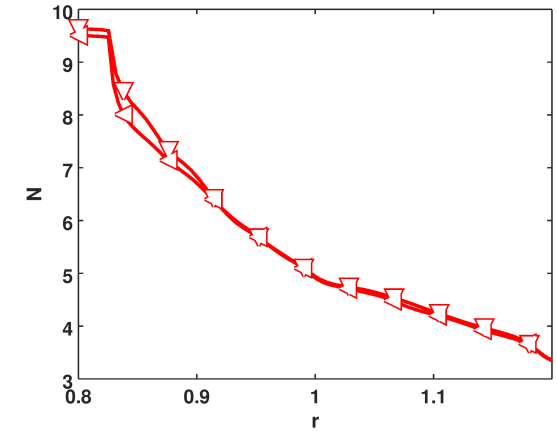
In this section we present general properties of turbulent solutions. A special attention is given to limiter configuration for which an extensive convergence study with respect to discretization has been performed. In both limited (Fig. 4a and Fig. 4b) and diverted (Fig. 4c and Fig. 4d) geometries, we recover characteristic features of turbulent regimes such as blobs of different sizes, exhibiting a clear and well known ballooning in the transport. Such turbulent systems can be described by decomposing the solution, in space and time, using the sum of an equilibrium plus a perturbation. In Sec. IV A, such equilibrium is studied while in Sec. IV B, high frequencies in space and time are studied.

A. Impact of the discretization on the self-organised plasma equilibrium

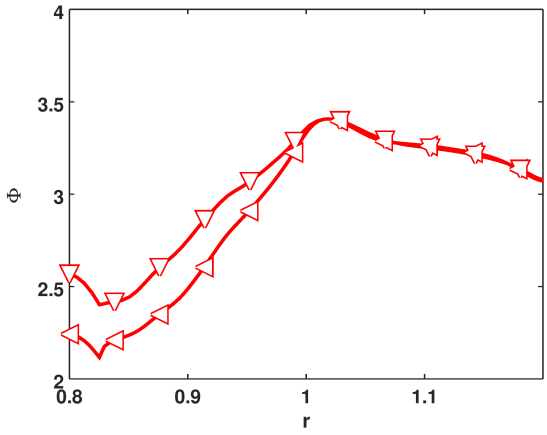
In this subsection we focus on the averaged quantities of the main unknowns, the density N, the electric field Φ, the vorticity W and the ionic parallel momentum Γ. Averages are performed over time for a duration of 50000 ω_c^{-1} and the two angles around the mid plane with $\theta \in [-\pi/4, +\pi/4]$, focusing on the low field side, and $\phi \in [0, \pi]$, describing half of the torus which is thus taken as π-periodic by definition. Only remains the radial dimension r, normalised to the minor radius. In order to estimate the consistency of any patterns, always two curves are shown, results of a simulation using a "regular" discretization ($64 \times 512 \times 32$) and a "refined" discretization ($96 \times 768 \times 48$), see Tab. III. The time span used for averaging is regarded as long enough not to be dazzled by fast and local reorganisation such as avalanches but is still subject to the slow dynamic of converging towards the equilibrium. It reveals density profiles with two slopes: one within and one outside the close fields lines area ($r < 1$) for both meshes (Fig. 1a). Two slopes are also present in the electric field profiles (Fig. 1b). While no discrepancy can be seen in the Scrape-Off Layer, a slightly different $\partial_r \Phi$ can be seen in the edge. This might lead to a slightly faster poloidal rotation but it did not seem to significantly impact the density profile as $\partial_r^2 \Phi$ are quite similar. Indeed, the vorticity $W = \vec{\nabla} \cdot (\frac{1}{B^2}(\vec{\nabla}_\perp \phi + \frac{1}{N}\vec{\nabla}_\perp N))$ is based on second derivatives of the density and electric fields, and both resolutions are in fair agreement (Fig. 1c). Finally, the parallel ionic momentum exhibits an offset. It is important to recall that we focus on an average performed in the most turbulent part of the poloidal section, $\theta \in [-\pi/4, \pi/4]$. It appears that an offset does not impact, for example, the contribution of the divergence of the the parallel ionic momentum along the field lines $\vec{\nabla} \cdot (\Gamma \vec{b})$ in the density equation, as seen in Fig. 1e. Nevertheless the equilibrium of Γ is quite sensitive as it is a balance between two very large contributions nearly cancelling each other plus some corrections. Any modifications of the mesh or the geometry leads to modification of the equilibrium. Fig. 1f shows the impact of increasing resolution on the $\Gamma = 0$ contour. High resolutions push the stagnation lines further away from the limiter, even crossing the equator on the low field side.

B. Dynamic of small scales

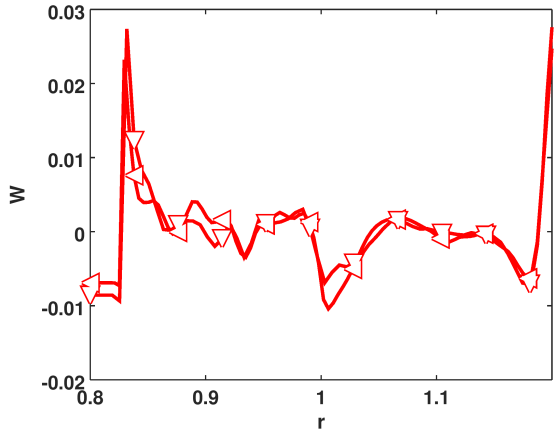
The intensity of the dynamic is easily quantified by calculating the $\mathcal{L}^2$-norm of $\{\partial_t \Psi\}_{ef}$, the effective time derivative of the solution simulated. Depending on the dimensions included in the norm and the dimension on which the norm is applied, this allows to recover spacial structure, time traces or even a single value. In particular one can easily test the condition Eq.(9) when comparing norms of $\{\partial_t \Psi\}_{ef}$ to norms of $\{\partial_t \Psi\}_{th}$. As



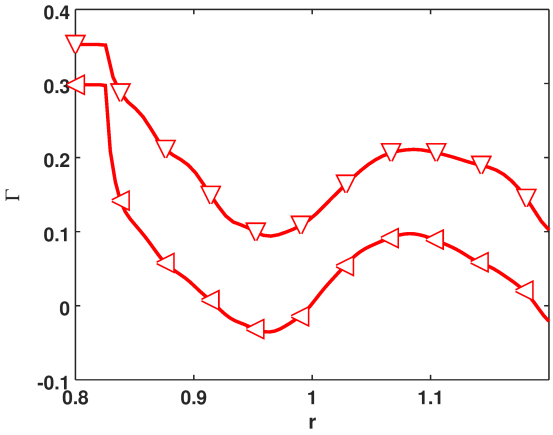
(a) Density N.



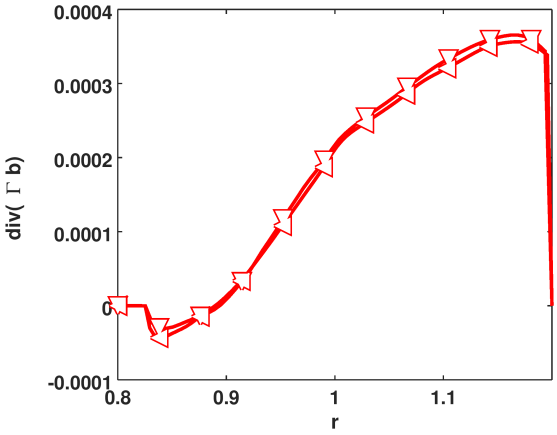
(b) Electric field Φ.



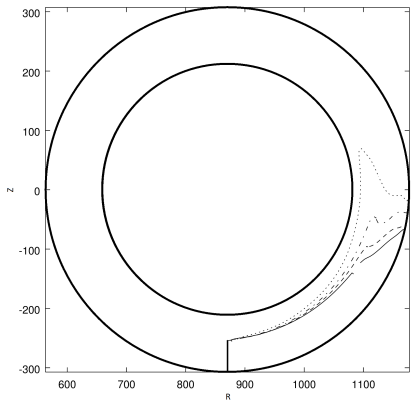
(c) Vorticity W.



(d) Parallel ionic momentum Γ.

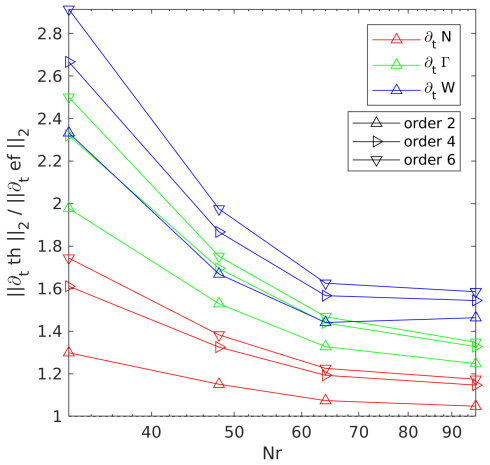


(e) $\vec{\nabla} \cdot \left(\Gamma \, \vec{b} \right)$

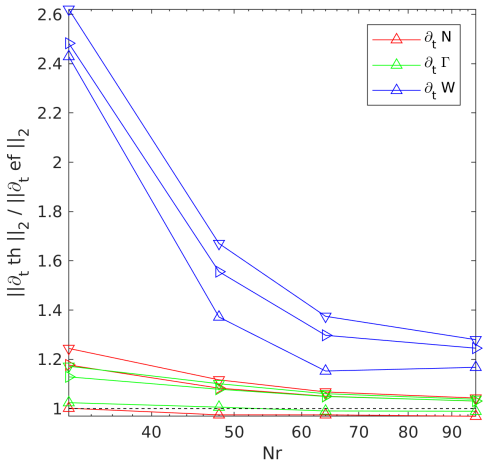


(f) $\Gamma = 0$ contour. $N_r = 32$ $-$, 48 - -, 64 $-\cdot$, 96 $\cdot\cdot$

FIG. 1: Plots of radial profiles of averaged quantities: N (a), Φ (b), W (c), Γ (d) and $\vec{\nabla} \cdot \left(\Gamma \, \vec{b} \right)$ (e). The averaging is carried out over $(t, \theta, \phi) \in ([0, 50000], [-\pi/4, \pi/4], [0, \pi])$. Regular mesh: $\triangledown$, refined mesh: $\triangleleft$; see Tab. III. In (f), 2D contours of $\Gamma = 0$ for $(t, \phi) \in ([0, 50000], [0, \pi])$. $D_N = 0.003$ and parameters given in Tab. IV.

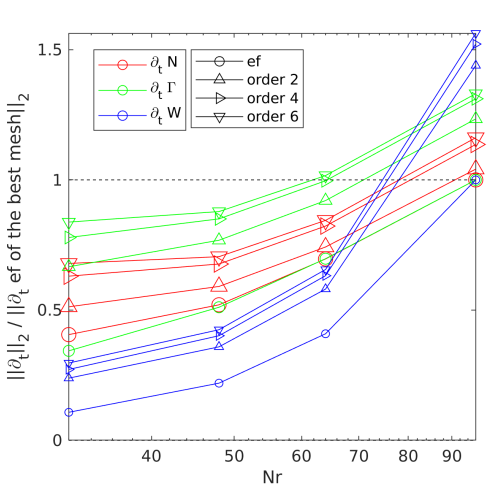


(a) Variation of the intensity of the full dynamic $||\partial_t X(r,\theta,\phi,t)||_2$ with N_r.

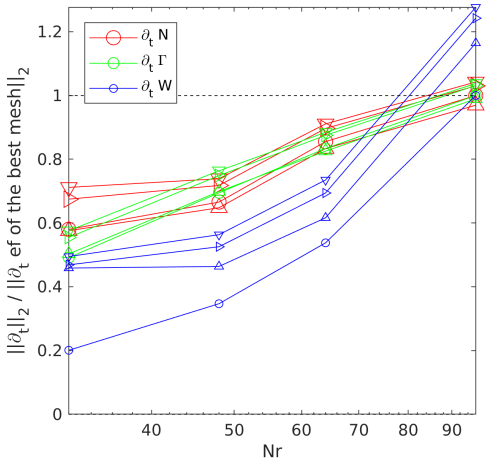


(b) Variation of the intensity of the mode-0 dynamic $|| < \partial_t X >_{\theta,\phi} (r,\theta,\phi,t)||_2$ with N_r.

FIG. 2: Plots of the ratio of the $\mathcal{L}_2$-norm of the off-line evaluations of $\{\partial_t N\}_{ol}$ (red), $\{\partial_t \Gamma\}_{ol}$ (green), $\{\partial_t W\}_{ol}$ (blue) over their respective effective measurement $\{\partial_t \times\}_{ef}$ as a function of N_r. Off-line computations based on second-order ($\triangle$), fourth-order ($\triangleright$), and sixth-order ($\triangledown$) approximations. Parameters Tab. IV excepted $D_N = 0.003$.



(a) Intensity of the full dynamic $||\partial_t X(r,\theta,\phi,t)||_2$.

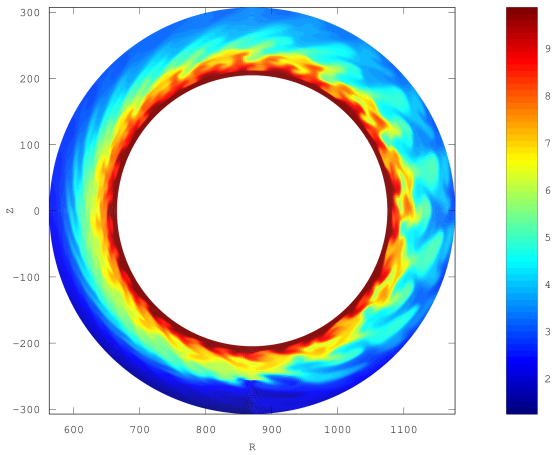


(b) Intensity of the mode-0 dynamic $|| < \partial_t X >_{\theta,\phi} (r,\theta,\phi,t)||_2$.

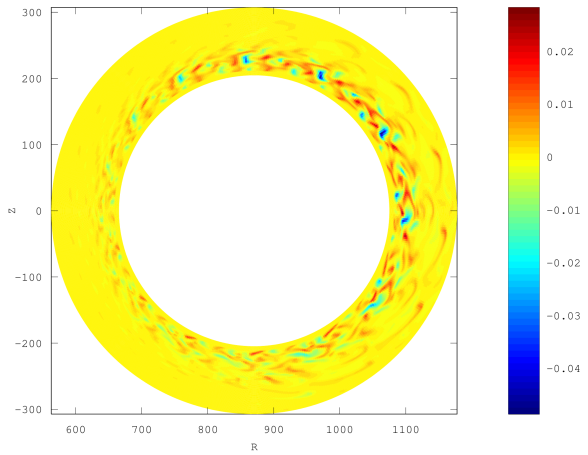
FIG. 3: Ratio of the $\mathcal{L}^2$-norm of off-line evaluations of $\{\partial_t N\}_{ol}$ (red), $\{\partial_t \Gamma\}_{ol}$ (red), $\{\partial_t W\}_{ol}$ (blue) divided by the respective effective measurement $\{\partial_t x\}_{ef}$ of the simulation using the best mesh. Off-line computations use second-order ($\triangle$), fourth-order ($\triangleright$), sixth-order ($\triangledown$) approximations. Effective measurements from TOKAM3X output uses circles ($\bigcirc$). Parameters Tab. IV excepted $D_N = 0.003$.

$\|\{O^i\}_{th}\|$ can not be computed, we simply study the convergence of $\|\{O^i\}_{ol}\|$ computed off-line with high order schemes.

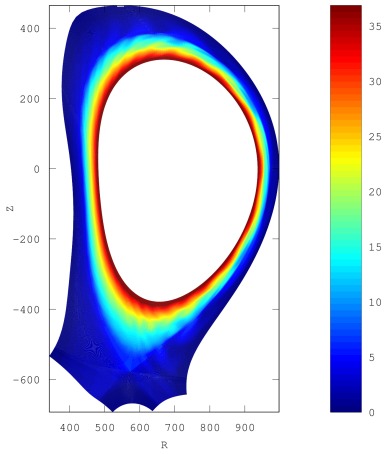
Fig. 2a shows the $\mathcal{L}^2$-norms of $\{\partial_t \Psi\}_{ol}$, $||\{\partial_t \Psi\}_{ol}(r,\theta,\phi,t)||^2_{r,\theta,\phi,t}$, for $\Psi = N, \Gamma, W$ depending on the resolution. The norms are calculated at orders 2, 4 and 6, and are normalized to the $\mathcal{L}^2$-norm of $\{\partial_t \Psi\}_{ef}$. At each order, and for each variable, the ratio decreases when increasing the resolution, thus $\{\partial_t \Psi\}_{ef}$ converges towards $\{\partial_t \Psi\}_{ol}$. Comparing the norms $\{\partial_t \Psi\}_{ol}$ at different orders, one can see the gap between sixth-order and fourth-order decreases with respect to the improvement of the mesh resolution while being significantly smaller than the gap between fourth-order and second-order. For such a reason, we consider the sixth-order computation to be close enough to the errorless theoretical computation in order to do



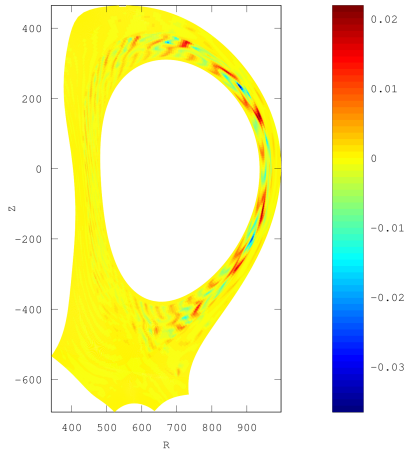
(a) Density.



(b) Time derivative of the density.



(c) Density.



(d) Time derivative of the density.

FIG. 4: Instantaneous $2D$ plots of density and its time derivative in turbulent regime and in arbitrary units. (a, b) limited geometry, (c, d) diverted geometry. $D_N = 0.003$ and parameters given in Tab. IV.

the approximation $\{O^i\}_{th} \simeq \{O^i\}_{ol}$. Increasing the order of the off-line measurement would not significantly change the results of this study. Further, Fig. 2a shows that using a sixth-order off-line evaluation leads to a stronger dynamics $\times 1.3$ to $\times 1.7$ than the one computed by TOKAM3X. This difference is not related to a problem of conservation as TOKAM3X is based on a conservative numerical scheme. The source driving the system is then totally transported towards the edges by the radial flux. $||\partial_t X||_2$ should not be confused with the intensity of radial flux which has a direction as $||\partial_t X||_2$ is more relevant of the direction less buoyancy of the system, buoyancy which increases when smaller smaller scales are described.

Similarly to Fig. 2a, Fig. 2b shows the $\mathcal{L}^2$-norms of the mode-0 of $\{\partial_t \Psi\}_{ol}$ defined by $< \{\partial_t \Psi\}_{ol} >_{\theta,\phi}$, then $|| < \{\partial_t \Psi\}_{ol} >_{\theta,\phi} (r, \theta, \phi, t)||^2_{r,\theta,\phi,t}$, for $\Psi = N, \Gamma, W$ and depending on the resolution. As the mode-0 is the main contribution to the equilibrium, it is interesting to see the norms of the reconstructed dynamics at order six is only overestimated by a factor $\times 1.1$ to $\times 1.4$. This means that most of the discrepancy between the theoretical and effective dynamics comes from small scales in space and time.

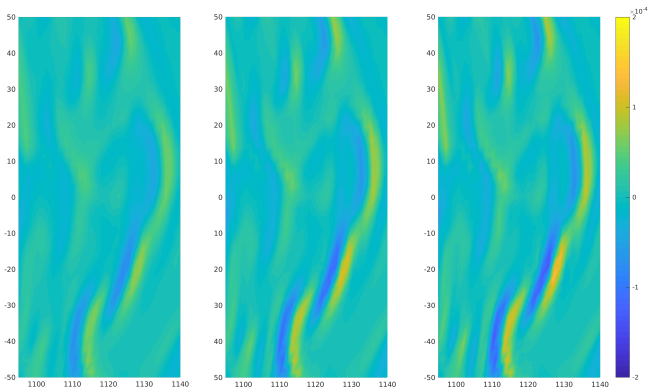
In the second series of plot, Fig. 3, the norm of the theoretical dynamic computed at different order is not normalized with respect to the effective dynamic at the considered mesh resolution but it is normalized with respect to the effective dynamic using the finest mesh (see Tab. III). Also, the effective dynamic is plotted. On top of the fact that increasing the mesh resolution increases the concordance between effective and theoretical dynamics, these series of plots show the absolute value of the norm 2 also increases with respect

to the mesh resolution. The focus on the mode 0 reveals the same qualitative behavior while quantitatively, the effect is significantly smaller.

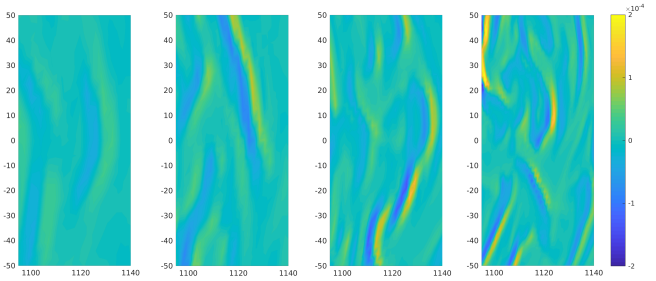
The highlighted discrepancies at small scales are better shown using $2D$ plots in Fig. 5. The diffusion operator of the density equation is displayed when computed off-line at different orders (Fig. 5a using a unique simulation on a regular mesh (see Tab. III)). Also, using a sixth-order off-line scheme, four simulations using different meshes are compared (Fig. 5b). On the one hand, increasing the order of the numerical scheme increases the intensity of the operator which can be much more peacked locally. On the other hand, increasing the mesh resolution also increases the intensity of the operator but it also changes the size and number of structures. In fact, simply using a simulation computed on a given mesh and then recomputing the density operator at different orders does not show any new self consistent organisation of the solution since the solution is not recomputed but only diagnosed. On the contrary, comparing different simulations using the same numerical scheme (TOKAM3X second-order) but different meshes allows to access new self organised regimes. In the present case, the new self organised regimes include more structures of smaller size.

This last study can be displayed using probability density function of the effective dynamic. Fig. 6 allows to quantify the strong deviation for rare / intense events with respect to the mesh.

In each sensitivity analysis of this section, it has been shown that the equilibrium (or mode 0) of the solution can be considered as fairly resolved while the small scales tend to be underestimated with respect to the number of small structures and their intensity. Also, the error on the density equation is smaller than on the parallel ionic momentum, itself lower than the error of the vorticity equation. Knowing that the spectrum of the vorticity is flatter than the one of the parallel momentum, itself flatter than the one of the density, we also draw the conclusion that the essential part of the error comes from the small scales. The simulations under investigation would benefit of an higher-order discretization scheme and/or an increased mesh resolution. In any case, the total radial transport flux would be the same as TOKAM3X is based on a conservative numerical scheme and a transport driven by a source. Still, the small scale self-organisation could be different, using a different number of structures of different size, with edges of different shape and steepness, also moving at different velocities or in different directions. Over the parameters used, no evidence of a possible qualitative regime change has been seen, only quantitative changes. Still, as turbulence is intrinsically non-linear and populates small scales, extrapolation of results is a tricky task.



(a) second-order (left), fourth-order (center), and sixth-order (right).



(b) $32 \times 256 \times 16$ (left), $48 \times 384 \times 24$ (left-center), $64 \times 512 \times 32$ (right-center), and $96 \times 768 \times 48$ (right).

FIG. 5: 2D plots of the diffusion operator of the density N computed in post-processing. (a) Impact of the discretization scheme order for a $64 \times 512 \times 32$ mesh. (b) Impact of the mesh size for a sixth-order numerical scheme. Zoom on the low field side in the poloidal plane (R, Z).

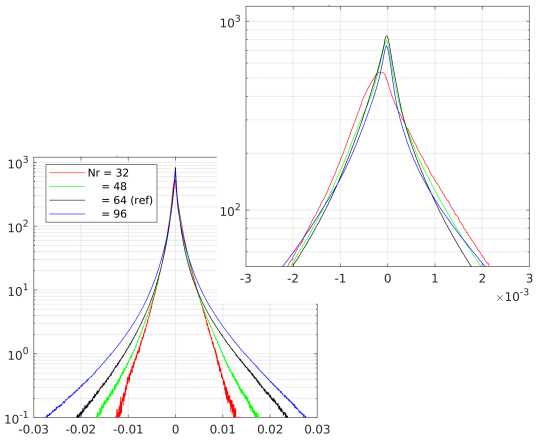
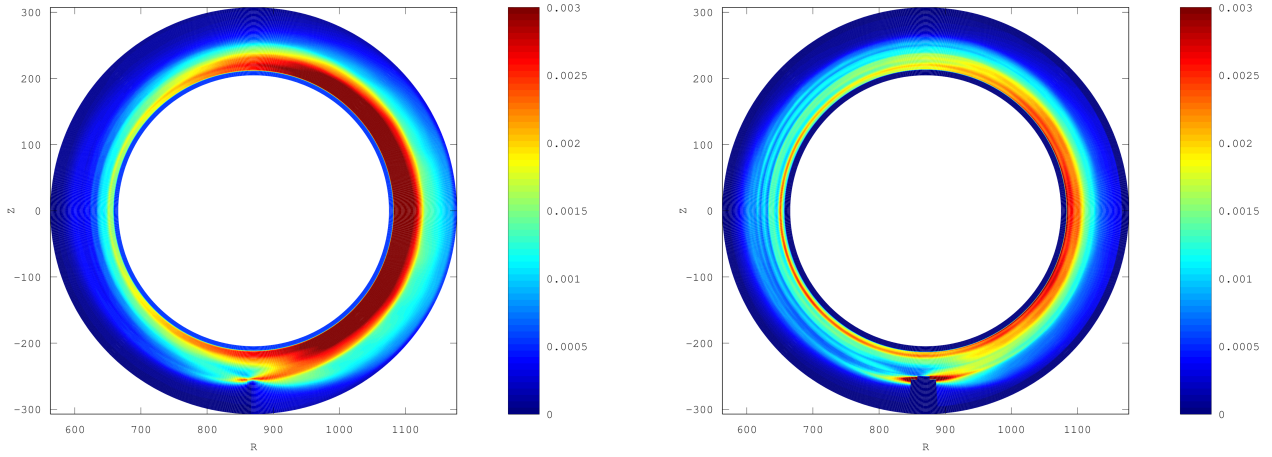
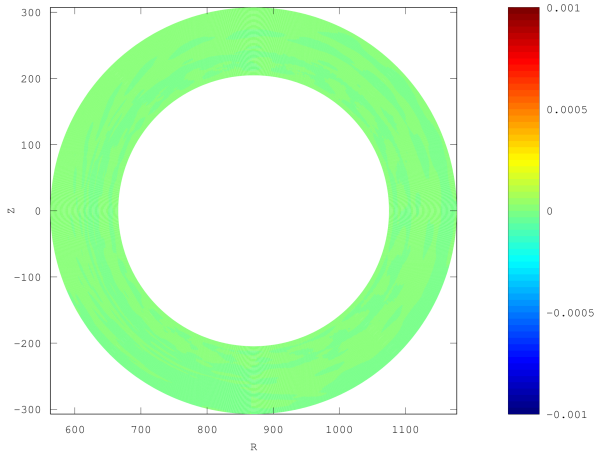


FIG. 6: Probability density function of $\{\partial_t N\}_{ef}$ measured from four simulations of TOKAM3X using different meshes (see Tab. III). Right figure focuses on $\{\partial_t N\}_{ef} = 0$.

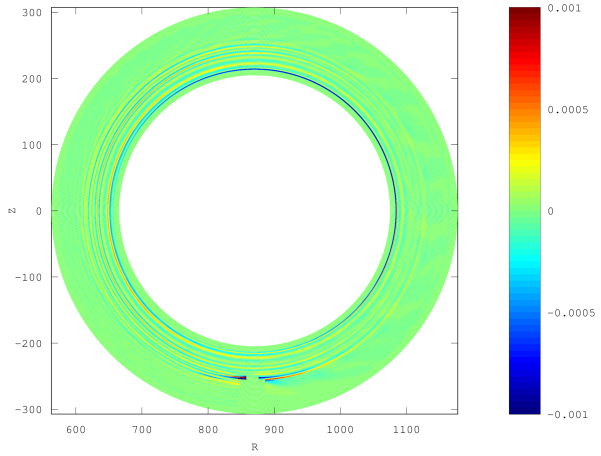


(a) Time and ϕ averages of $|\{\partial_t N\}_{ef}|$

(b) Time and ϕ averages of $|\epsilon_{to}|$



(c) Time and ϕ averages of $\{\partial_t N\}_{ef}$



(d) Time and ϕ averages of ϵ_{to}

FIG. 7: $2D$ plots of averaged quantities: Fig. 7a $= |\partial_t N|$, Fig. 7b $= |\epsilon_{to}|$, Fig. 7c $= \partial_t N$ and Fig. 7d $= \epsilon_{to}$. The averaging is performed for $(t, \phi) \in ([0, 50000], [0, \pi])$. Limiter geometry, $D_N = 0.003$ and parameters given in Tab. IV

V. ERROR QUANTIFICATION OF TWO TYPICAL 3D TURBULENT SIMULATIONS IN DIVERTOR AND LIMITER CONFIGURATIONS

In order to interpret (i)PoPe results, one has to look first at residuals (Sec. V A) and then at weights (Sec. V B). Indeed, if a too large residual is measured, then weights would be of a secondary importance because a large residual means that a large part of the dynamic is not captured by the operators used for the projection. Both residuals and weights are computed using projections of $3D$ subsets of points in (r, θ, ϕ), points taken at a given time t. It induces a time dependence for the weights of operators and also for any norm of residuals. In Tab. I, a sense of the variability in the time series is given by the first and last ninth deciles.

A. Residuals structure and statistics

To understand and then decide if a residual is acceptable, one can use either the total residual $\epsilon_{to} = \{\partial_t X\}_{ef} - \{\partial_t X\}_{th}$ or the effective residual ϵ_{ef} in order to compare it to to either the theoretical dynamic $\{\partial_t X\}_{th}$ or the effective dynamic $\{\partial_t X\}_{ef}$. In this study, as we already emphasized, the ratio between $||\{\partial_t X\}_{ef}||_2$ and $||\{\partial_t X\}_{th}||_2$ can go up to $\times 1.7$ on a regular mesh. Here we focus on the ratio $||\epsilon_{to}||_2 / ||\{\partial_t X\}_{ef}||_2$ which also contains information about $\{\partial_t X\}_{th}$ by the definition of ϵ_{to}. Tab. II contains the mean value, first and ninth deciles over time of the ratio of each of the four equations. The equation on the electric field Φ shows an error of less than 20% in both geometry as well as a relative narrow distribution of the ratio around its mean value (very close

first and ninth deciles). Evolution equations of the density, parallel ionic momentum and vorticity contain significantly higher discrepancy, without any surprise compared to the $\times 1.7$ ratio $||\{\partial_t X\}_{th}||_2/||\{\partial_t X\}_{ef}||_2$. With respect to the study of the ratio, we recover here a similar hierarchy of errors, the vorticity being the most impacted by the numerical error, then the parallel momentum, and last the density. The equation determining the electric field not including any time derivative, it discards an important source of error while still not being exempted of errors.

Residual and dynamic lay in a $4D$ space (r, θ, ϕ, t) and can be studied with a deeper insight than just some statistics of a ratio. Fig. 7 reveals $2D$ structures in the poloidal plane of the density equation for a limiter geometry. Fig. 7a points out that the activity of the dynamics $|\{\partial_t N\}_{ef}|$ is mostly located in the edge and in particular on the low field side where the transport is more important. Directly comparing with Fig. 7c where we remove the absolute value of the diagnosis, computing simply $\{\partial_t N\}_{ef}$, we draw the conclusion that the simulation is at equilibrium as $< \{\partial_t N\}_{ef} >_{\phi,t} \simeq 0$ and thus the dynamic seen in Fig. 7a is not due to any equilibrium still building in the simulation. The norm of the total error measured by $|\epsilon_{to}|$ (Fig. 7b) contains a distribution similar to $|\{\partial_t N\}_{ef}|$ plus a peak of error at the top of the limiter and in its vicinity. Indeed, the top of the limiter is crossed by strong variations of the different fields. Finally it is also important to measure that on average the contribution of errors tends to zero, $< \epsilon_{to} >_{\phi,t} \simeq 0$ in Fig. 7d.

Even if residuals are significant, no extreme value have been found already knowing the difficulties of simulating small scales. A study of each operator's weight in each equation is then meaningful.

B. Weights statistics

The amplitude of weights are also given using 3 values: the first decile, the mean and the ninth decile. The W equation which requires the inversion of a Laplacian is quite well solved, both in limiter and divertor configuration. The three other equations, $\partial_t N$, $\partial_t \Gamma$ and $\partial_t W$ have quite similar behaviour. In particular the diffusion operator has a weight always higher than the theoretical value, suggesting some numerical diffusion. According to the considered mesh and plasma parameters, the divertor configuration seems to be closer to the theoretical values. The $E \times B$ contribution has a value nearly two time smaller than the theoretical one while the operator for the parallel transport is very close to the theoretical value.

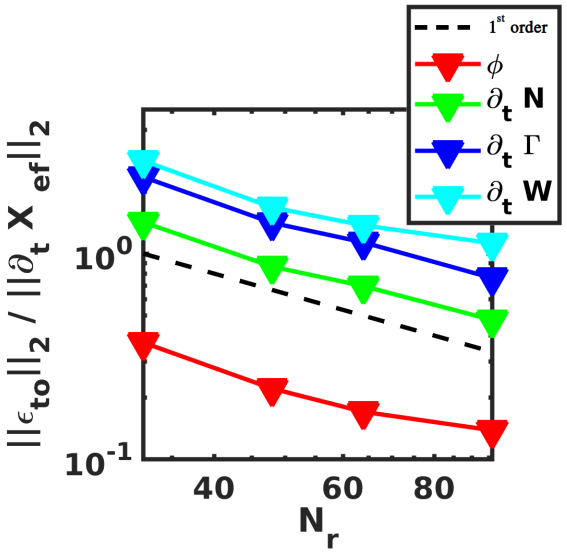


FIG. 8: Ratio of $||\epsilon_{to}||_2/||\{\partial_t X\}_{ef}||_2$ for each equation, with respect to N_r using four different meshes (Tab. III) in a limiter configuration.

VI. ACTUAL ORDER OF THE NUMERICAL SCHEME

Four meshes have been used to recover the actual order of the numerical scheme when simulating production simulations, see Tab. III. The "regular" mesh has been used for a large number of simulations in different studies. Such mesh allows the version of TOKAM3X used for this study to run efficiently across many core/nodes while requiring a reasonable amount of RAM. Indeed, TOKAM3X uses an implicit solver based on the inversion of a large ill conditioned $3D$ matrix. The "refined" mesh is defined using a scaling factor $\times 6/4$ with respect to the number of points in each dimension, $\times 6/4$ being the maximum we could offer ourselves for such a study. The "coarse" mesh is not the smallest one allowing for qualitative simulations but it is a configuration which have been significantly used in the past.

Fig. 8 the ratio of $||\epsilon_{to}||_2/||\{\partial_t X\}_{ef}||_2$ is displayed with respect to the mesh resolution for each of the equations. While TOKAM3X is based on a second-order WENO scheme, we only recover a first-order convergence. This point is due to the fact that the WENO scheme can locally downgrade its approximation to a first-order when encountering large gradients. In addition to this, as seen in Sec. IV, the solution has a constrained spectrum because of the resolutions limitations. This second point implies that changing the resolution one also slightly changes the regime of the simulation, going to more turbulent dynamics. Having more turbulent dynamics then tends to increase the numerical error and to degrade the order of the scheme just below 1.

Equation	Operator	Limiter	divertor
$W =$	$\vec{\nabla}\cdot\left((\vec{\nabla}_\perp\Phi)/B^2\right)$	0.954 : 0.957 : 0.961	0.963 : 0.967 : 0.972
	$\vec{\nabla}\cdot\left((\vec{\nabla}_\perp \ln N)/B^2\right)$	0.964 : 0.974 : 0.984	0.988 : 0.996 : 1.004
$\partial_t N =$	$\vec{\nabla}\cdot\left(D_N\vec{\nabla}_\perp N\right)$	3.741 : 4.448 : 5.169	1.589 : 1.711 : 1.852
	$-\vec{\nabla}\cdot\left(\Gamma\ \vec{b}\right)$	0.975 : 1.005 : 1.037	1.093 : 1.149 : 1.204
	$-\vec{\nabla}\cdot(N\vec{u}_E)$	0.570 : 0.681 : 0.799	0.528 : 0.579 : 0.626
	$-\vec{\nabla}\cdot\left(N\vec{u}^{ion}_{\nabla B}\right)$	0.816 : 0.915 : 1.017	0.997 : 1.009 : 1.021
$\partial_t \Gamma =$	$\vec{\nabla}\cdot\left(D_\Gamma\vec{\nabla}_\perp\Gamma\right)$	3.039 : 3.551 : 4.163	1.776 : 1.885 : 1.994
	$-\vec{\nabla}\cdot\left(\Gamma^2/N\vec{b}\right)$	1.023 : 1.046 : 1.070	1.041 : 1.088 : 1.140
	$-\vec{\nabla}\cdot(\Gamma\vec{u}_E)$	0.492 : 0.574 : 0.672	0.480 : 0.526 : 0.575
	$-\vec{\nabla}\cdot\left(\Gamma\vec{u}^{ion}_{\nabla B}\right)$	0.929 : 1.015 : 1.102	0.901 : 0.941 : 0.975
	$-\nabla_\parallel(2N)$	0.816 : 0.882 : 0.935	0.250 : 0.299 : 0.356
$\partial_t W =$	$\vec{\nabla}\cdot\left(D_W\vec{\nabla}_\perp W\right)$	2.181 : 2.521 : 2.906	1.404 : 1.522 : 1.624
	$-\vec{\nabla}\cdot\left(\Gamma W/N\vec{b}\right)$	1.296 : 1.731 : 2.156	1.598 : 2.137 : 2.625
	$-\vec{\nabla}\cdot(W\vec{u}_E)$	0.606 : 0.664 : 0.728	0.581 : 0.631 : 0.688
	$2\vec{\nabla}\cdot\left(N\vec{u}^{ion}_{\nabla B}\right)$	0.915 : 0.937 : 0.954	0.995 : 0.996 : 0.998
	$1/\eta_\parallel\vec{\nabla}\cdot\left[\left(\nabla_\parallel log\,(N)-\nabla_\parallel\Phi\right)\vec{b}\right]$	0.779 : 0.829 : 0.876	0.897 : 0.919 : 0.940

TABLE I: Table summarizing the model equations together with their respective weights for "reference simulations" described in Tab. IV except $D_N = 0.003$. Weights are given with the value of the first decile, the mean and the ninth decile. The source S_N that should be present in the equation "$\partial_t N$" is not taken into account as the verification is performed on a subsets of points where this source is strictly null.

.

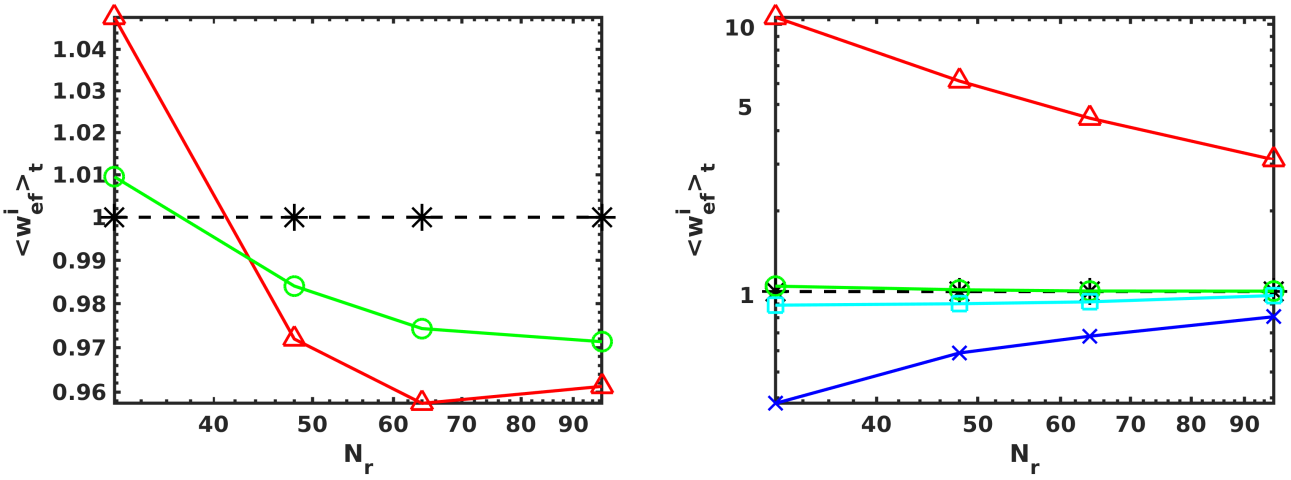
Equation	Limiter	divertor
W	0.16 : 0.17 : 0.19	0.17 : 0.18 : 0.20
$\partial_t N$	0.50 : 0.70 : 0.89	0.83 : 0.95 : 1.08
$\partial_t \Gamma$	0.84 : 1.14 : 1.44	1.52 : 1.73 : 1.94
$\partial_t W$	1.04 : 1.37 : 1.76	1.79 : 2.19 : 2.60

TABLE II: Table of the ratio between the total residual ϵ_{to} and the off-line estimation $\{\partial_t X\}_{ol}$ computed at the sixth-order For the W equation, the right hand side W is taken as the theoretical reference. $D_N = 0.003$ for the limiter geometry, otherwise parameters given in Tab. IV. Measurement are given with the value of the first decile, the mean and the ninth decile.

	N_r	N_θ	N_Φ	number of points	scaling to ref
coarse	32	256	16	131072	2/4
medium	48	384	24	442368	3/4
regular	64	512	32	1048576	4/4 ref
refined	96	768	48	3538944	6/4

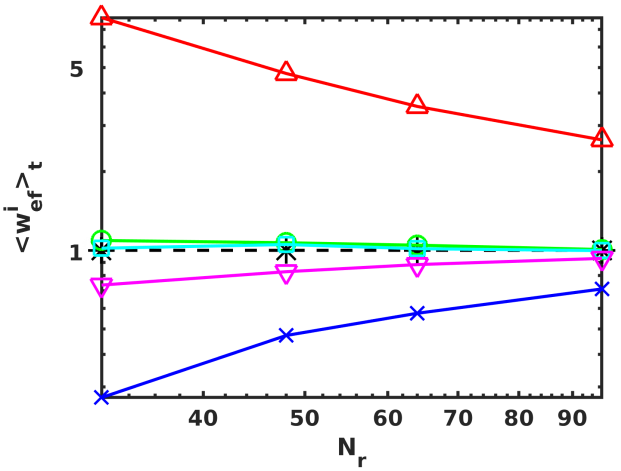
TABLE III: Resolutions for meshes used in the limiter geometry. The same ratio is kept between r, θ and ϕ dimensions for the different meshes.

Convergence can also be qualitatively measured looking at the weight of each operator in each equation, Fig. 9a, Fig. 9b, Fig. 9c and Fig. 9d. Each plot has been constructed using four simulations in limiter geometry, $D_N = 0.003$ and parameters given in Tab. IV. Often, like for the equation of the electric field (Fig. 9a), weights clearly converge but towards a value slightly off with respect to the theoretical value. For this particular equation, similarly to the residual study, we recover a good agreement with the theoretical value of weights. The den-

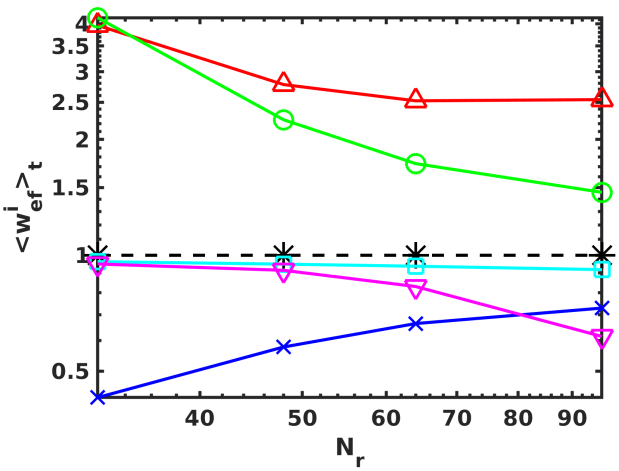


(a) W equation's weights with respect to N_r.

(b) $\partial_t N$ equation's weights with respect to N_r.



(c) $\partial_t \Gamma$ equation's weights with respect to N_r.



(d) $\partial_t W$ equation's weights with respect to N_r.

FIG. 9: The color legend is based on the following sequence of operator listed in Tab. I: red up triangles = 1st, green circles = 2sd, blue crosses = 3rd, cyan squares = 4th, magenta down triangles = 5th. Four meshes are used in a limiter configuration, see Tab. III. $D_N = 0.003$, other parameters Tab. IV.

sity equation (Fig. 9b) and the parallel momentum equations both behave according to the same trends (Fig. 9c): the diffusion operator (red) is overestimated by a factor 4 for a regular mesh while the divergence of the $E \times B$ (blue) is underestimated by 30%. Weights of the vorticity equation's operators are mostly converging and stay within a narrower range than the density and parallel momentum except the last operator, the divergence of the parallel current (magenta) which is an extremely steep operator.

VII. STUDY OF THE NUMERICAL DIFFUSION

In this section, a sensitivity analysis is conducted by varying the parameter D_N, the diffusion coefficient of the density equation. As in the previous section, we use four different mesh resolutions detailed in Tab. III to better understand trends. This coefficient has a strong impact on the solution as it controls the spectrum of the density. Decreasing D_N increases the amplitude of high frequency modes corresponding to small structures, thus flatten the spectrum. As the density is coupled to the ionic parallel momentum, the vorticity and the electric field, controlling the spectrum of the density allows to control the spectrum of each system's unknowns. It is important to recall that for any value of D_N, TOKAM3X is able to perform a simulation with almost perfect conservation as it is based on a conservative scheme and as long as the time step of the explicit part of the numerical scheme satisfies the CFL condition. Nevertheless, we point out that requesting a given D_N in the input of a computation code does not insure that the output will indeed embed the theoretical value.

Fig. 10a is the key figure of this section as it expresses the effective weight of D_N with respect to its theoretical weight. One can clearly see a saturation effect, a plateau defining a minimal value for D_N effective, starting at around $D_N = 10^{-2}$ for the mesh prescribed. As seen on Fig. 9b, one can see on Fig. 10a that the saturation level of the plateau depend on the mesh resolution, a better discretization leads to a lower plateau. This is easily understood as a high resolution mesh can better describe the dynamics of smaller structures. In this way, this plot allows to estimate a numerical diffusion threshold with respect to the numerical scheme and the discretization. Such analysis also expresses that the simulations using $D_N \in [0.0003, 0.003]$ should not differ too much as the effective D_N is almost constant. Prior to study other graphs, it is worth to notice for the highest value studied, $D_N = 0.5$, the system is not turbulent. This is why the value of operators' weight does not follow the same trend as for the range $D_N \in [0.0003, 0.05]$. Indeed, weights of other operators can also be influenced by the value of D_N because the spectrum of each unknown depend on D_N. For the density equation, the $E \times B$ advection term would be the second most impacted one, systematically underestimated, see Fig. 10c. The WENO scheme used to discretize this non linear term tends to smooth the intensity of the $E \times B$ contribution, not by underestimating the overall radial transport as this radial transport is always respected by such a conservative scheme, but locally, by decreasing the mixing effect of blobs. Once again, the saturation level depend on the discretization used: the mesh with the highest number of degree of freedom handles better the non-linear dynamic considered as its weight is closer to the theoretical value of 1. Also the saturation effect appears at lower values of D_N for higher resolutions. Operator $-\vec{\nabla} \cdot (\Gamma\ \vec{b})$ (Fig. 10b) and operator $-\vec{\nabla} \cdot (N \vec{u}^{ion}_{\nabla B})$ (Fig. 10d) of the density equation appear to be much less sensitive. Still the value of weights in the range of $D_N \in [0.0003, 0.003]$ tends be constant, emphasising the similarities between any simulations using $D_N \leq 0.003$.

As the model strongly couples each variable, the value of D_N is likely to also impact operators' weight of other equations by impacting the spectrum of other unknowns. Indeed, the diffusive operator $\vec{\nabla} \cdot (D_\Gamma \vec{\nabla}_\perp \Gamma)$ of the momentum equation also tends to be over represented compared to the theoretical value, see Fig. 11a. Similarly to $-\vec{\nabla} \cdot (N \vec{u}_E)$ in $\partial_t N$, the operator $-\vec{\nabla} \cdot (\Gamma \vec{u}_E)$ in $\partial_t \Gamma$ struggles to reach the theoretical amplitude of 1, see Fig. 11b. Nevertheless, the vorticity equation determining the electric field is not significantly impacted, see Fig. 11c and Fig. 11d. In all four cases, we also recover trends with respect the mesh resolution and plateaus in the range of $D_N \in [0.0003, 0.003]$.

Similar trends are also observed in diverted geometry. Both a numerical diffusion (for $\vec{\nabla} \cdot (D_N \vec{\nabla}_\perp N)$ in the density equation (Fig. 12a) and for $\vec{\nabla} \cdot (D_\Gamma \vec{\nabla}_\perp \Gamma)$ in the moment equation (Fig. 12c)) as well as an underestimation of the $E \times B$ contribution (in the density equation (Fig. 12b) and the momentum equation (Fig. 12d)) are present in the form of plateaus.

As a conclusion, numerical methods, mesh and state of the system matter to estimate the error.

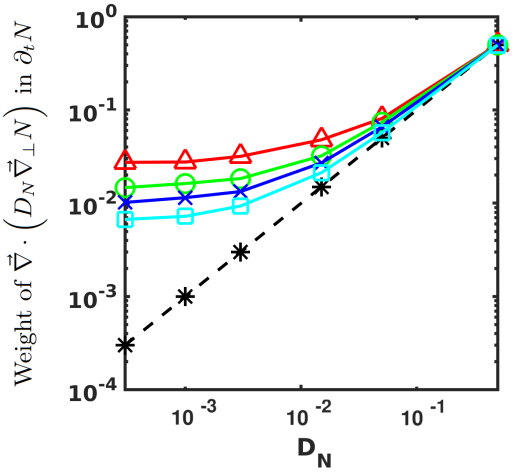
VIII. SUMMARY

This paper introduces iPoPe, a data-mining method derived from PoPe. Using the output of a simulation, both methods allow to do *a posteriori* error estimate by recovering the equations effectively used in a simulation. Their direct results are the effective value of the weights corresponding to the various terms prescribed in the mathematical model plus an additional error term called the residual.

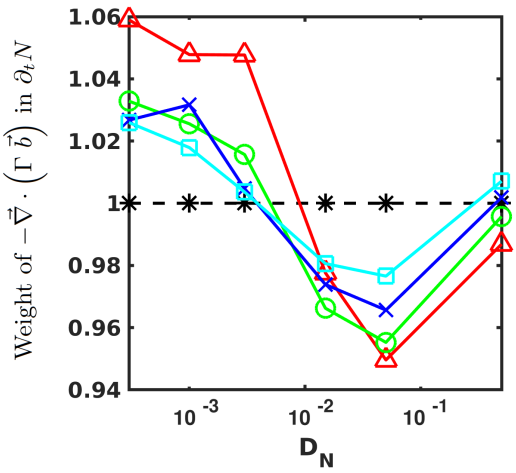
iPoPe method is successfully applied here to fluid simulations of edge plasma in tokamak. It shows its capability to address realistic geometries and boundary conditions, as well as any flow regime including turbulence, by analysing the state-of-the-art 3D code TOKAM3X. It extends in that sense the verification usually performed by the Method of the Manufactured Solutions while not requiring any dedicated simulations but merely a computational overhead of 1% for each simulation verified. However, the challenging part is to interpret (i)PoPe results. Indeed measuring the discrepancy between the theoretical and the effective values of the operators weights is easier than predicting the impact of this discrepancy on the solution. In some cases, predicting the behavior of the error could be as challenging as predicting the behavior of the system without the error.

Besides the zero-error case, the ideal situation is when the numerical error can be expressed analytically. If it can be expressed analytically, this expression can be found using (i)PoPe by adding new operators in the projection step. The determination of those new operators should be guided by the properties of their weights (average and standard deviation with respect to time) and the norm of the total residual. Two cases are thus possible:

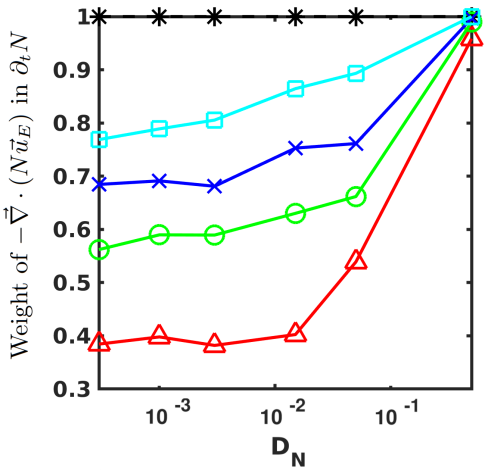
- The analytical expression obtained for the error does not contradict the basic assumptions driving the model and neither changes the nature of the dynamics. In this case, the error can be considered as being part of the theoretical model, and it becomes acceptable. It is the case when the error is related to numerical diffusion or to the use of a flux limiter when it remains low or well-separated with respect to the space and time scales of interest. It could be seen as diffusive subgrid models taking into account a physics at a scale not present in the theoretical model.



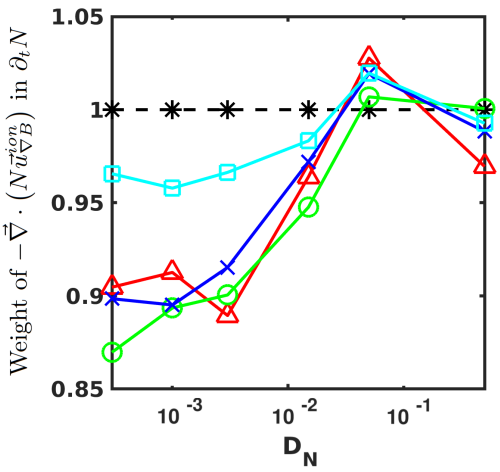
(a) Weight of $\vec{\nabla}\cdot\left(D_N\vec{\nabla}_\perp N\right)$ in $\partial_t N$.



(b) Weight of $-\vec{\nabla}\cdot\left(\Gamma\,\vec{b}\right)$ in $\partial_t N$.



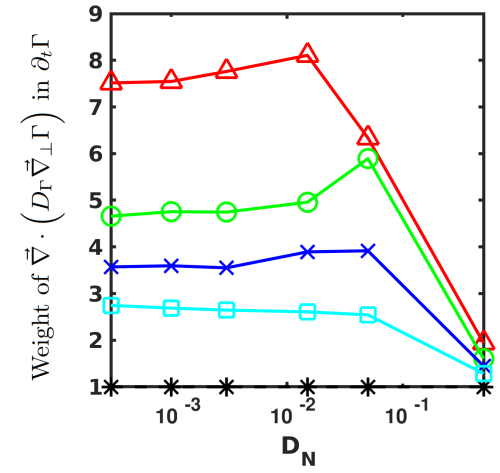
(c) Weight of $-\vec{\nabla}\cdot\left(N\vec{u}_E\right)$ in $\partial_t N$.



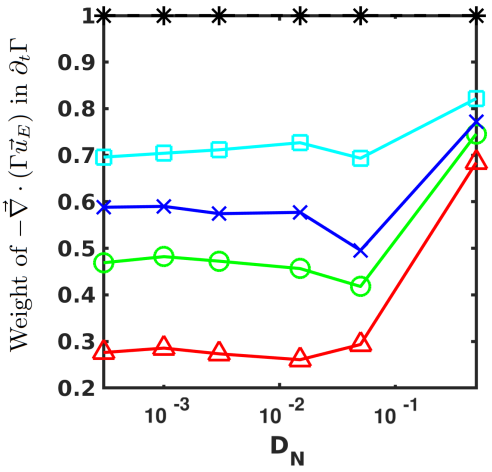
(d) Weight of $-\vec{\nabla}\cdot\left(N\vec{u}^{ion}_{\nabla B}\right)$ in $\partial_t N$.

FIG. 10: Sensitivity analysis of different operators' weight when variating D_N, the diffusion coefficient of the density equation. Four mesh resolutions are used: $32\times256\times16$ red triangles ($\triangle$), $48\times384\times24$ green circles ($\circ$), $64\times512\times32$ blue crosses ($\times$) and $96\times768\times48$ cyan squares ($\square$). Limiter geometry, other parameters Tab. IV.

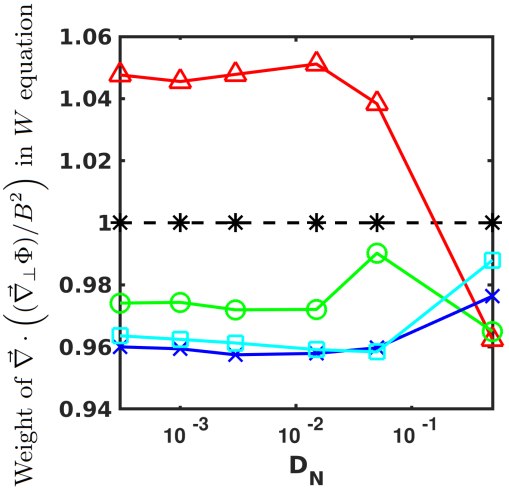
- The error is not acceptable because no analytical expression could capture it, or it is in too strong interaction with the theoretical dynamics of the system. Moreover, it cannot be minimised at an acceptable cost. In this case second case, (i)PoPe provides a measurement of such artefact and it could be used to guide the methodological choices by answering the following questions: which is the most appropriate discretization approach? how to allocate the degrees of freedom in space and time? which algorithm could minimize the error? An operator can be isolated in the splitting, and then solved using a similar discretization but with an higher order, or a different discretization like a pseudo-spectral approach. Another solution can be to periodically filter the error, similarly to a dealiasing filter used in pseudo-spectral algorithms. It is worth to mention a study[23] of growth rates depending on a narrow resonance where the numerical error is large (100% or $(i)PoPe = 0$) but somehow orthogonal to the process studied, and thus with a low impact on the features of the simulated solution.



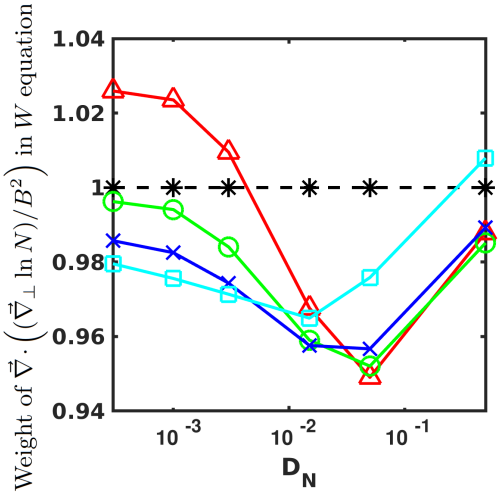
(a) Weight of $\vec{\nabla}\cdot\left(D_\Gamma\vec{\nabla}_\perp\Gamma\right)$ in $\partial_t\Gamma$.



(b) Weight of $-\vec{\nabla}\cdot(\Gamma\vec{u}_E)$ in $\partial_t\Gamma$.

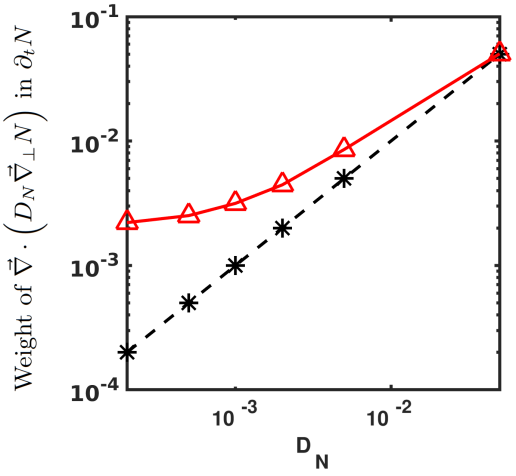


(c) Weight of $\vec{\nabla}\cdot\left((\vec{\nabla}_\perp\Phi)/B^2\right)$ in W equation.

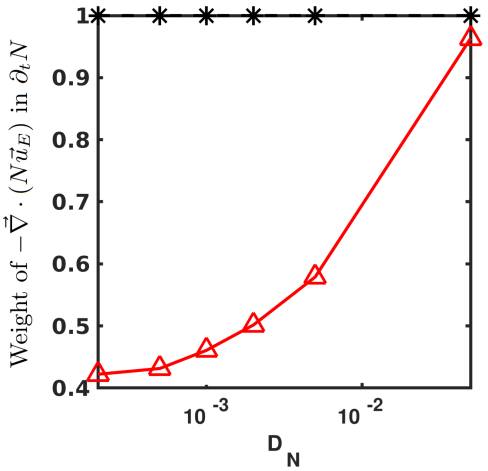


(d) Weight of $\vec{\nabla}\cdot\left((\vec{\nabla}_\perp\ln N)/B^2\right)$ in W equation.

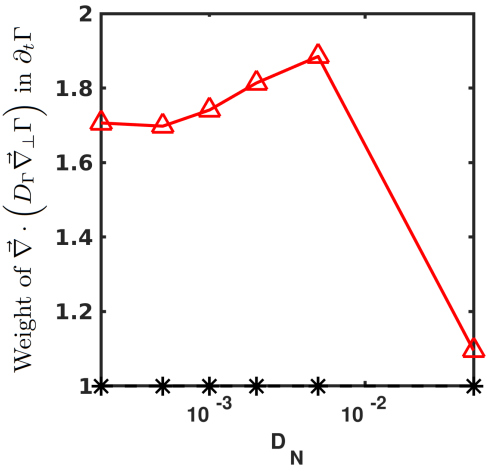
FIG. 11: Sensitivity analysis of different operators' weight when variating D_N, the diffusion coefficient of the density equation. Four mesh resolutions are used: $32\times256\times16$ red triangles ($\triangle$), $48\times384\times24$ green circles ($\circ$), $64\times512\times32$ blue crosses ($\times$) and $96\times768\times48$ cyan squares ($\square$). Limiter geometry, other parameters Tab. IV.



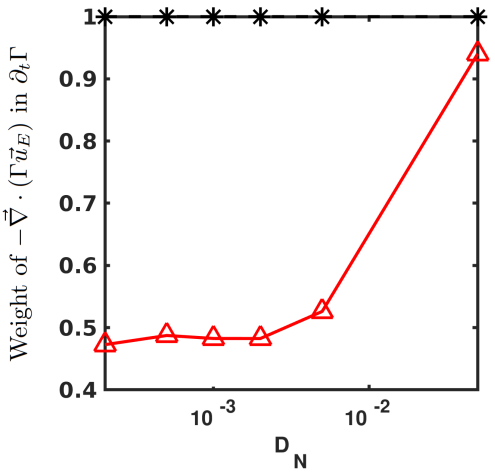
(a) Weight of $\vec{\nabla}\cdot\left(D_N\vec{\nabla}_\perp N\right)$ in $\partial_t N$.



(b) Weight of $-\vec{\nabla}\cdot(N\vec{u}_E)$ in $\partial_t N$.



(c) Weight of $\vec{\nabla}\cdot\left(D_\Gamma\vec{\nabla}_\perp\Gamma\right)$ in $\partial_t\Gamma$.



(d) Weight of $-\vec{\nabla}\cdot(\Gamma\vec{u}_E)$ in $\partial_t\Gamma$.

FIG. 12: Sensitivity analysis of different operators' weight when varying D_N, the diffusion coefficient of the density equation. Diverted geometry, other parameters Tab. IV.

ACKNOWLEDGEMENTS

This work was granted access to the HPC resources of IDRIS, under the allocations i2018056912 made by GENCI, and of Aix-Marseille University, financed by the project Equip@Meso (ANR-10-EQPX-29-01). The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University-A*MIDEX, a French "Investissements d'Avenir programme". This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 20142018 under grant agreement No. 633053 for the project WP17ENRCEA08.

Appendix A: Data used for simulations

Geometry	Limiter	Divertor
A	3.4	2.8
$\rho^\star = \rho_L/a$	1/256	1/256
L_φ	$\pi/2$	$\pi/2$
$\eta_\parallel (B_0/en_0)$	10^{-5}	10^{-5}
D_N , D_Γ , $D_W(\rho_L^2\omega_C)$	$5\cdot10^{-3}$	$5\cdot10^{-3}$
$S_N^0(\rho_L^{-3}\omega_C)$	$4\cdot10^{-3}$	$4\cdot10^{-3}$
N_ψ (edge)	32	40
N_θ (edge)	512	350
N_ψ (SOL)	32	40
N_θ (SOL)	512	350
N_ψ (divertor leg)		32
N_θ (divertor leg)		16
N_ψ (PFR)		9
N_θ (PFR)		16
N_φ	32	32

TABLE IV: Table of the reference simulations with associated parameters.

[1] T. Cartier-Michaud, P. Ghendrih, Y. Sarazin, J. Abiteboul, H. Bufferand, G. Dif-Pradalier, X. Garbet, V. Grandgirard, G. Latu, C. Norscini, C. Passeron, P. Tamain; Projection on Proper elements for code control: Verification, numerical convergence, and reduced models. Application to plasma turbulence simulations; Physics of Plasmas 2016, 23 - 2 020702

[2] K. Salari, P. Knupp; Code verification by the method of manufactured solution; Sandia report SAND2000-1444 (2000), http://prod.sandia.gov/techlib/access-control.cgi/2000/001444.pdf

[3] P. J. Roache; Code Verification by the Method of Manufactured Solutions; Journal of Fluids Engineering, 2001 nov - 1, 4–10 - 124

[4] P. Knupp, K. Salari; Verification of Computer Codes in Computational Science and Engineering; 2002

[5] W. L. Oberkampf, T. G. Trucano; Verification and validation in computational fluid dynamics; Prog. Aerosp. Sci. 2002, 38-209

[6] P. Tamain, Ph. Ghendrih, H. Bufferand, G. Ciraolo, C. Colin, N. Fedorczak, N. Nace, F. Schwander, E. Serre; Multi-scale self-organisation of edge plasma turbulent transport in 3D global simulations; Plasma Physics and Controlled Fusion, 2015 - 5, 054014 - 57

[7] F. Riva, P. Ricci, F. D. Halpern, S. Jolliet, J. Loizu, A. Mosetto; Verification methodology for plasma simulations and application to a scrape-off layer turbulence code; Physics of Plasmas 2014, 6 - 21

[8] C. J. Roy; Review of code and solution verification procedures for computational simulation; Journal of Computational Physics, 2005 - 1, 131 - 156 - 205

[9] T. Cartier-Michaud; Vérification de Codes et Réduction de Modèles : Application au Transport dans les Plasmas Turbulents; PhD manuscript, 2016; https://tel.archives-ouvertes.fr/tel-01252592

[10] C. S. Chang, S. Ku; Particle Simulation of Neoclassical Transport in the Plasma Edge; Contributions to Plasma Physics, 2006, 46, 79, 496-503

[11] C.S. Chang, S. Ku, G.R. Tynan, R. Hager, R.M. Churchill, I. Cziegler, M. Greenwald, A.E. Hubbard, J.W. Hughes; Fast Low-to-High Confinement Mode Bifurcation Dynamics in a Tokamak Edge Plasma Gyrokinetic Simulation; Phys. Rev. Lett., 118, 17, 175001, 6, 2017

[12] S. I. Braginskii; Transport processes in a plasma; Reviews in Plasma Physics, Vol.1, Ed. M.A. Leontovich (New York Consultant Bureau), (1965) 205-311

[13] A. N. Simakov, P. J. Catto; Drift-ordered equations for modelling collisional edge plasma; Contr. Plasma Phys. 44(1-3) (2004) 83-94

[14] B.D. Dudsona, M.V. Umanskyb, X.Q.Xub, P.B.Snyderc, H.R. Wilson; BOUT++: A framework for parallel plasma fluid simulations; Computer Physics Communications, 180, 9, 1467-1480, 2009

[15] P Ricci, F D Halpern, S Jolliet, J Loizu, A Mosetto, A Fasoli, I Furno, C Theiler; Simulation of plasma turbulence in scrape-off layer conditions: the GBS code, simulation results and code validation; Plasma Phys. Control. Fusion, 54 (2012) 124047

[16] V. Naulin, T. Windisch, O. Grulke; Three-dimensional global fluid simulations of cylindricalmagnetized plasmas, Phys. of Plasma, 15 (2008) 012307

[17] P. Tamaina, H. Bufferand, G. Ciraolo, C. Colin, D. Galassi, Ph. Ghendrih, F. Schwander, E.Serre; The TOKAM3X code for edge turbulence fluid simulations of tokamak plasmas in versatile magnetic geometries; Journal of Computational Physics, 321, 606 - 623, 2016

[18] F. Riva, C. Colin, J. Denis, L. Easy, I. Furno, J. Madsen, F. Militello, V. Naulin, A. H. Nielsen, J. M. B. Olsen, J. T. Omotani, J. J. Rasmussen, P. Ricci, E. Serre, P. Tamain, C. Theiler; Blob dynamics in the TORPEX experiment: a multi-code validation; Plasma Phys. Control. Fusion 2016

[19] C. Colin, P. Tamain, F. Schwander, E. Serre, H. Bufferand, G. Ciraolo, N. Fedorczak, Ph. Ghendrih; Impact of the plasma-wall contact position on edge turbulent transport and poloidal asymmetries in 3D global turbulence simulations; Journal of Nuclear Materials, 463, 654 - 658, 2015

[20] D. Galassi, P. Tamain, H. Bufferand, G. Ciraolo, Ph. Ghendrih, C. Baudoin, C. Colin, N. Fedorczak, N. Nace, E. Serre; Drive of parallel flows by turbulence and large-scale E × B transverse transport in diverted geometry; Nuclear Fusion, 57, 3, 036029, 2017

[21] P. C. Stangeby; The plasma boundary of magnetic fusion devices; IOP (2000)

[22] http://pastix.gforge.inria.fr/

[23] T. Cartier-Michaud, Ph. Ghendrih, G. Dif-Pradalier, X. Garbet, V. Grandgirard, G. Latu, Y. Sarazin, F. Schwander, E. Serre; Verification of turbulent simulations using PoPe: quantifying model precision and numerical error with data mining of simulation output; Journal of Physics: Conference Series, 2018, 1125-012005